RESULTS OF AN ANALYSIS OF SDSS GALAXIES IN THE VO

W. Schoenell[1], **M. Cerviño**[2,3], **R. Cid Fernandes**[1], **A. Mateus**[4], **E. Terlevich**[5], **R. Terlevich**[5], **F. de los Santos**[5], **J.P. Torres-Papaqui**[1,5], **and V. Luridiana**[2,3]

[1]Universidade Federal de Santa Catarina (UFSC; Brazil)
[2]Spanish Virtual Observatory (SVO; Spain)
[3]Instituto de Astrofísica de Andalucía (IAA-CSIC; Spain)
[4]Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG-USP; Brazil)
[5]Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE; Mexico)

ABSTRACT

We present here the VO access to the results of an analysis of the spectra of Sloan Digital Sky Survey (SDSS) galaxies performed with the STARLIGHT code by Cid Fernandes et al. (2005). The results include for each galaxy the original SDSS spectrum, the best-fit synthetic spectrum, the star formation history, the pure emission line spectrum corrected for the underlying stellar population (in SDSS emission line galaxies) and the intensity of several emission/absorption lines. The database will be accessible from the PGos3 server and it will be released at the end of summer 2007.

Key words: Stellar populations; Virtual Observatory.

1. INTRODUCTION

The interoperability concept promoted by the Virtual Observatory (VO) is not only limited to data sharing, but it also includes enhancing the possibility to analyze these data (usually by means of comparison with theoretical models) and to infer relevant physical quantities. This task can only be performed if front-end VO interfaces (mostly, applications) are able to provide access to theoretical data (see Rodrigo's et al. contribution in these proceedings) and if analysis tools are implemented in a VO compliant fashion. Theoretical and observational data must not be hardwired in the VO-analysis tools: rather, they should be obtained from VO services. This is the only way to minimize the model-dependent bias in the analyzed data.

The use of VO services to input theoretical and observational data into VO analysis tools creates diversity and improves productivity: since the VO provides a homogeneous structure for data sharing (VOTables), it allows to design automatic tools and workflows that can use different sets of observational and theoretical data without making program drivers for each specific data set. Once data providers distribute their results in VOTables, they can use these analysis tools for further research.

Although this is in principle the most desirable scenario, there are several difficulties that make this kind of workflow difficult to establish, mainly: (i) There is no user-friendly way to recover the credits of the data used in current VO applications; an issue mentioned several times during this meeting and especially relevant for theoretical studies. (ii) The access and exploration of VO-enabled services with theoretical datasets[1] are not implemented in most VO-applications, with VOSpec and VOSED as the only exceptions.

In this contribution we describe the first stages of the implementation of the analysis tool STARLIGHT (Cid Fernandes et al., 2005) in the VO framework.

2. THE STARLIGHT CODE

The STARLIGHT code is designed to obtain the best fit to an observed spectrum, O_λ, taking into account the corresponding error σ_{obs}, with a theoretical model spectrum, M_λ, using a Markov Chains Monte-Carlo algorithm with simulated annealing. The code finds the minimum χ^2,

$$\chi^2 = \sum_\lambda \left(\frac{O_\lambda - M_\lambda}{\sigma_{\mathrm{obs}}} \right)^2 , \qquad (1)$$

---

[1]Examples are the PEGASE-HR (see Ph. Prugniel's contribution in these proceedings) and PGos3 databases (see Cerviño's et al. contribution in these proceedings) for evolutionary synthesis models at `http://vo.obspm.fr/cgi-bin/siap/pegasehr.pl` and `http://ov.inaoep.mx/pgos3` respectively.

and obtains the corresponding physical parameters of the modeled spectrum: (i) the star formation history, x_j, as a function of a base of N_{SSP} Single Stellar Population (SSP) models normalized at λ_0, $b_{j,\lambda}$ (ii) the extinction coefficient of predefined extinction laws, r_λ, and (iii) the velocity dispersion σ_v, which obey the relation:

$$M_\lambda = M_{\lambda_0} \left(\sum_{j=1}^{N_{\text{SSP}}} x_j \, b_{j,\lambda} \, r_\lambda \right) \otimes G(v_*, \sigma_*). \quad (2)$$

With this technique, we have obtained the physical parameters of the galaxies of the SDSS data release 2 (York et al., 2000; Stoughton et al., 2002; Abazajian et al., 2003, 2004), using the library of SSP models by Bruzual & Charlot (2003) as database. Since the modeled spectrum provides the stellar continuum of the galaxy, we can also measure emission line intensities and equivalent widths. This huge computational effort resulted in a unique database with more than half a million galaxies, which occupies close to 500 GB of data.

3. VO IMPLEMENTATION

The VO implementation of STARLIGHT allows to explore different issues in the VO, from describing the results of the data analysis (including references to input data, models and processes) to establishing a VO workflow using theoretical VO services.

We created a preliminary web service to access and manipulate this database and obtain the corresponding VOTables at `http://www.starlight.ufsc.br/`. Users can perform queries on data by SQL commands or spatial selections. The final VO service is expected to be available at the servers `http://www.starlight.ufsc.br/` and `http://ov.inaoep.mx`.

Currently we are working on two different aspects:

1. To produce tools to explore the database in a VO compliant way instead of a Web service. First essays using TSAP (Theoretical Spectral Access Protocol, see Rodrigo's et al. talk in these proceeding for extensions to non spectroscopic data) are not satisfactory enough due to the database complexity. At this point, a more general, recursive, protocol is needed.

2. To allow the STARLIGHT code to obtain SSP databases from (non-local) VO services. We are making this implementation using SSP models included in PGos3 database (see Cerviño's et al. contribution in these proceedings) as reference.

However, other issues remain to be addressed. Maybe the most important ones are:

1. How to manage references to VOTables nested in VOTables: the description of the star formation history of a galaxy needs nested references to VOTables with both original data and the SSP database used. At the same time, the SSP database used would need references to the assumed isochrones and atmosphere models, and so on.

2. How to know the coverage, uncertainties and range of validity of different SSP models in an automatic way. This issue can be partially addressed by the use of the current proposed IVOA recommendations on data model characterization and the spectral data model (see Cerviño & Luridiana contribution in these proceedings for their possible applications to SSP models).

REFERENCES

Abazajian, K., et al. 2004, AJ, 128, 502

Abazajian, K., et al. 2003, AJ, 126, 2081

Bruzual, G. & Charlot, S. 2003, MNRAS 344, 1000

Cid Fernandes, R., et al. 2005, MNRAS, 358, 363

Stoughton, C., et al. 2002, AJ, 123, 485

York, D. G., et al. 2000, AJ, 120, 1579